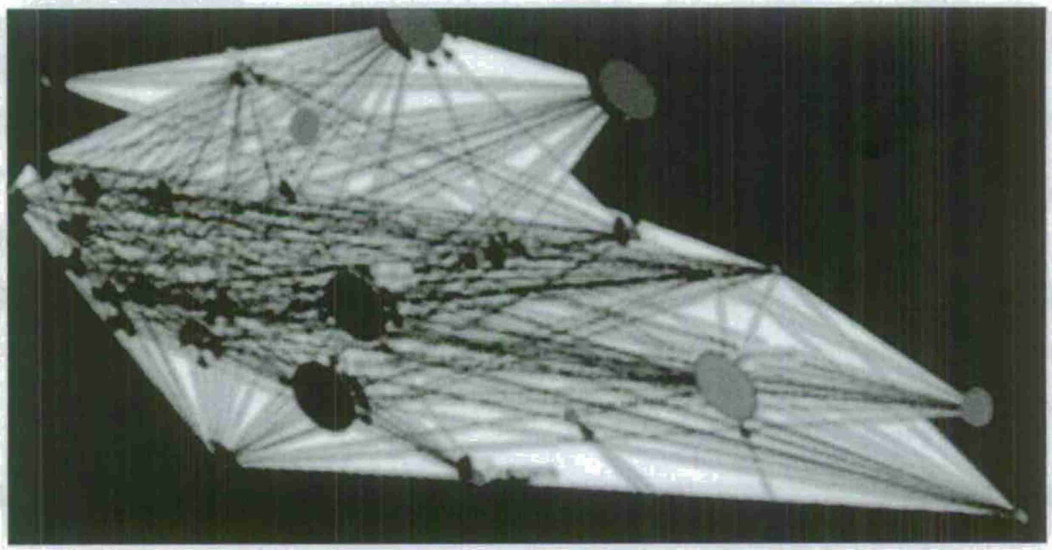
Computational Models for Belief Revision, Group Decision-Making and Cultural Shifts
MURI FINAL REPORT 25 Oct 2010



Abstract: This MURI (1) explores how beliefs support and lead to certain actions in one culture but not another, and (2) develops computational models that further our understanding of the relation between beliefs, decisions, and actions. One key requirement for such models is to distinguish the different roles played by instrumental or secular values versus sacred values, which differ widely across cultures. The goal of these models is to provide formal explanations for how the beliefs of individuals affect group and individual actions, and how groups with shared interests evolve. Such models are an important step toward understanding and predicting the dynamics and actions of groups. They are fundamental to an understanding of how actions of a group may be altered by belief revision, by either internal or external pressures (including force). They are also needed for strategic reasoning in negotiations, where beliefs in different cultures may lead to what appears to be irrational proposals, yet are seen as rational in that culture.

Worthwhile models require relevant data. In many cases, especially regarding the roles of sacred values and network structures, data were very sparse or not available. Significant efforts were conducted in the early years to fill this need. Consequently some of our models are still quite preliminary, but will help pave the way for future studies.

AFOSR MURI Grant No. A9550-05-1-0321 Dr. Terence Lyons, Program Mgr.
Principal Investigator: Whitman Richards, Massachusetts Institute of Technology

Computational Models for Belief Revision, Group Decision-Making and Cultural Shifts

MURI FINAL REPORT Oct 2010

AFOSR Grant No. A9550-05-1-0321 (Start Date: 1 May 05)
PRINCIPAL INVESTIGATOR: Professor Whitman Richards, 617-253-5776, wrichards@mit.edu
LEAD INSTITUTION: Massachusetts Institute of Technology
AFOSR PROGRAM MANAGER: Dr. Terence Lyons, 703-696-9542, terence.lyons@afosr.af.mil

Preface

Seven years ago, John Tangney, AFOSR, held a meeting in McLean VA to explore interdisciplinary approaches to understanding cultural differences in adversarial decision-making. The disciplines represented included Anthropology, Cognitive Science, Computer Science, Psychology, and Economics/Game Theory. This meeting was the birth of a novel thrust into a difficult and frustrating problem. Its solution required bridges between Mathematical approaches to the Behavioral Sciences as well as experimental studies over a wide spectrum of the Social Sciences. The MIT MURI was one of the efforts charged with exploring this intersection. The (final) group of MURI participants included ten universities and twelve disciplines (see below). The first challenge was for members of the MURI to understand better the approaches and mind set of the others, and to explore collaborative projects. This step took two years, and led to insights that any individual researcher working alone would have missed or ignored. At the end of five years, a total of over 120 publications document our efforts (see Section V.) This report provides brief summaries of the more important findings: (1) those relevant to DoD interests, and (2) advances in bringing more formal computational models into the soft underbelly of the Social Sciences. Although this second product of the MURI will take longer to assimilate in the scientific community, several of our experimental and field studies have had immediate impact, as witnessed by briefings to many Government agencies, Intelligence committees, and even the House of Lords in Britain.

Thanks go to Dr. Terence Lyons, AFOSR, for his generous support, and also for his ability, with Stephanie Bruce, to smooth over some bumps in the funding. The scope of the Lyons program also helped generate new ideas, as well as providing a competitive but friendly atmosphere for encouraging new initiatives beyond the original proposal. We are also indebted to Robert Axelrod for guidance on directions of our research. Also, special thanks to the many graduate students who participated in this program and helped knit

the collaborators together. At MIT, Alissa Cardone was superb in her administrative help with various aspects of the MURI, especially with the "Belief Dynamics" website, as well as the organization of several of our annual meetings. At participating universities, Andrea Fatica, (John Jay) and Jennie Woodring (Northwestern) also made life easier for us all. For readers interested in more details about the MURI research, the four annual reports give bullets and references of our findings each year, and can be found at http://groups.csail.mit.edu/belief-dynamics.

Whitman Richards 25 Oct 2010
MIT Computer Science & Artificial Intelligence Lab
Cambridge, MA 02139

Program Objective:

The primary objective of this MURI was (1) to explore how beliefs support and lead to certain actions in one culture but not another, and (2) to develop computational models that further our understanding of the relation between beliefs, decisions, and actions. One key requirement for such models is to distinguish the different roles played by instrumental or secular values versus sacred values, which differ widely across cultures. The goal of these models was to provide formal explanations for how the beliefs of individuals affect group and individual actions, and how groups evolve. Such models are an important step toward understanding and predicting the dynamics and actions of groups. They are fundamental to an understanding of how actions of a group may be altered by belief revision, by either internal or external pressures (including force). They are also needed for strategic reasoning in negotiations, where beliefs in different cultures may lead to what appears to be irrational proposals, yet are seen as rational in that culture.

Worthwhile models required relevant data. In many cases, especially regarding the roles of sacred values and network structures, data were very sparse or not available. Significant efforts were conducted to fill this need, although again, yet more work needs to be done. Consequently some of our models are still quite preliminary, but do help pave the way for future studies.

MURI Consortium Members *:

- **Massachusetts Institute of Technology** (PI) Whitman Richards, Prof. of Cognitive Science, Computer Science and Artificial Intelligence Laboratory (CSAIL),
  Joshua Tenenbaum, Assoc. Prof. of Computational Cognitive Science,
  Patrick Winston, Prof. of Computer Science. CSAIL
- ***Harvard University,** Avi Pfeffer, Assoc. Prof. of Computer Science
- **University of Michigan,** Scott Page, Prof. of Economics & Complex Systems,
  Jenna Bednar, Assoc. Prof. of Political Science
- **Northwestern University,** Kenneth Forbus, Prof. of Computer Science,
  Douglas Medin, Prof. of Psychology
- **CUNY, John Jay Center for Terrorism & Univ. Mich.,** Scott Atran, Prof. of Anthropology, Psychology and Public Policy
- **Georgia Tech,** Jeff Shamma,, Prof. Electrical & Computer Engineering
- **North Carolina State,** Jon Doyle, Institute Prof. of Computer Science
- **Georgetown,** John Mikhail, Assoc. Prof. Law & Philosophy (Ethics)
- **ARTIS Research,** R. Davis, CEO

Consultants: Robert Axelrod, Prof. of Political Science and Public Policy, University of Michigan; Marc Sageman, MD, PhD; Sageman Consulting LLC; Univ. Penn. Adjunct Prof., Dept. Psychiatry; Rajesh Kasturirangan, PhD., National Inst. Adv. Studies, Bangalore; J. J. Koenderink, Tech Univ Delft (formerly Prof. of Physics, Helmholtz Institute, Utrecht)

* Currently at Charles River Analytics, Cambridge, Ma.

Overview: This report is divided into four parts: I. Experimental Observations; II. Model Frameworks (including some simulations) III. Formal Models, IV. Transitions and V. Publications.

Understanding the origin and nature of shared and revised beliefs among group members in a culture is critical. Unfortunately, as mentioned, data are scarce, and hence an important component of our effort was the reliable documentation of the formation of groups such as those engaged in the acts at Madrid, London, Bali, etc. The research led by Scott Atran under our MURI currently provides one of the most detailed and reliable set of data available. This work augmented earlier studies of non-violent groups subject to cultural forces conducted by Atran & Medin (e.g. the Wisconsin Menomenee, the Guatemalan Ladinos, and Amish and Muslim communities.) Such studies of non-violent communities allowed closer examination of moral issues and sacred vs. secular values – the latter being critical in both belief revision and negotiations across cultures (See also

work by Bednar & Page on cultural signatures, and Tenenbaum regarding the relation among beliefs and community structure.) Because the origins of many beliefs rest in the traditions of a culture, Forbus, and Finlayson & Winston have formalized how beliefs, actions and values are expressed in stories within a culture. These formal frameworks for representing stories support rapid, automatic analysis of the semantic interpretation of text, such as excerpts from news sources. Implications for future trends can be explored through analogies with traditional stories from the relevant culture. In a related effort, Axelrod, Atran and Davis have laid out a theory for case-based influence. Finally, individuals or groups striving for social change need to have strategies for actions. Pfeffer & Gal have developed new frameworks for understanding strategic reasoning, subject to different beliefs, which dictate in part which moves are considered rational. These complement the more traditional "game-theoretic" approaches, and appear more plausible in many real world, multi-agent scenarios.

Beliefs underlie actions. These beliefs may be moral or cultural norms, perhaps the inferred threat of force or new incentives or rewards, or simply a sense of what one's peers expect of you. Closely associated with actions are assessments or evaluations of the risk and consequences of failure. Belief revision thus entails a complex interaction of cultural norms, risk, preference, reasoning, habits and peer relations. Doyle has provided a formal framework for belief change; Mikhail outlines models (based on legal codes) for moral behaviors in a culture. In addition to the several formal theories, the experimental side of the MURI has discovered new facts that bear on the nature of beliefs, reasoning, and actions when cultural differences come into play. We begin by highlighting some of these observations, and how they have revised our views and models.

I Experimental Observations

A. Cultural Clashes & Sacred Values

• The most important rule of thumb is that conflicts over sacred values (such as holy land) typically can not be resolved simply by monetary incentives. Indeed, offering monetary rewards to others who regard land (or a site) as sacred may backfire (Ginges et al 2007; Ginges & Atran, 2009; 2009; Bennis et al, 2010).

The above result has been documented in conflict situations in the Middle East, Southeast Asia, India, North Africa, Europe as well as here in North America (Native Indians and "urbanites"). Sacred values differ from material or instrumental values in that they incorporate moral beliefs that drive actions in ways dissociated from prospects for success. Across the world, people believe that devotion to essential or core values – such as the welfare of their family and country or their commitment to religion, honor, and justice – are, or ought to be, absolute and inviolable. A corollary of these studies is that cost-benefit analyses may be of little value in the resolution of conflicts involving sacred values, and offering material benefits actually makes agreements more difficult because people see the offering as an insult.

• Making symbolic concessions of no apparent material benefit often might open the way to resolving seemingly irresolvable conflicts.

On a more positive note, although much previous work on sacred values suggests they are absolute and intractable, more recent studies now show

Fig 1: If sacred issues are at stake, cost-benefit arguments for conflict resolution become irrelevant.

that some sacred values may be subject to contextual effects, or especially to a reframing of one's position to demonstrate respect for the other party's

values (Davis et al 2010). Especially relevant is the Iranian nuclear program (Dehghani et al, 2009). Unlike sacred land, the relevant sacred value is national sovereignty, with the perceived right to engage in a research program strengthened by external pressures. Here we have a sacred value put in place over a short time frame, as contrasted to land dispute that may go back hundreds of years. Given the nature of these short-term sacred value developments, there is a greater opportunity for flexibility.

• Although moral principles often play a central role in activating sacred values, their relevance need not make a value sacred. Thus we need to separate moral concerns from sacred perceptions. For example (see Bennis et al 2010.)

B. Moral principles

Acts within one culture, such as stoning, public hangings, or cannibalism may be seen as barbaric or extreme (e.g. suicide bombers) by members of one culture, but appropriate within another. These differences, including religious norms, affect negotiations when the cultures clash. Surprisingly, there has been no extensive study of which moral principals are "universal" and which vary considerably between cultures. A step in filling these lacunae has been taken by Mikhail (2010), with a study of acceptable homicides of 204 nations, as defined by the legal codes for a nation. The underlying hypothesis is that legal systems codify the acceptable <moral> behaviors for that nation-state. Note that legal systems evolve over the centuries, so in parallel we can infer developments in moral behaviors, which in turn may partially affect what is considered "sacred". (See comment above in previous section on absolute vs flexible sacred values.)

• The main finding is clear: throughout the world, intentional killing without justification or excuse is prohibited, and that self defense and insanity and to a lesser extent provocation is sometimes a valid justification. The opener here for change in behavior is "justification", which becomes the key factor sanctifying for terrorist acts.

In parallel, Mikhail has completed a study of 26 moral probes that underlie behaviors. This study also sets the stage for future behavioral research. Especially important is a potential tool for designing alternative routes to

sacred value negotiations is now available, using framing approaches to change perceived context and intent. In a later section (III-B), a formal model for evaluating moral dilemmas is proposed, one that can help guide framing attempts.

C. Terrorist Motivations

What motivates a terrorist? Without some insight into this question, productive steps can not be taken to combat violent acts. Large scale studies based on interviews. Here are two samples obtained by S. Atran:

• "Our people don't own airplanes and tanks, only human bombs. Those who carry out martyrdom operations are not retarded, not hopeless, not poor, but are the best of our people. They do not flee from life. They are educated, not illiterate, successful in their lives" (Sheikh Hamed al-Betawi, 2004).

" There is no nobler life than to die as a martyr for Jihad. None. The highest deed in Islam is Jihad. If we commit to Jihad, we can neglect other deeds, even fasting and prayer." (Abu Bakr Ba"asyir, 2005)

The key is to understand that in many cases, moral beliefs alone can drive some actions (e.g. terrorist or suicide bombing), which implies that certain acts are dissociated from prospects for success. These are difficult to countermand (see qualifier in section I-A.) For example, devotion to certain

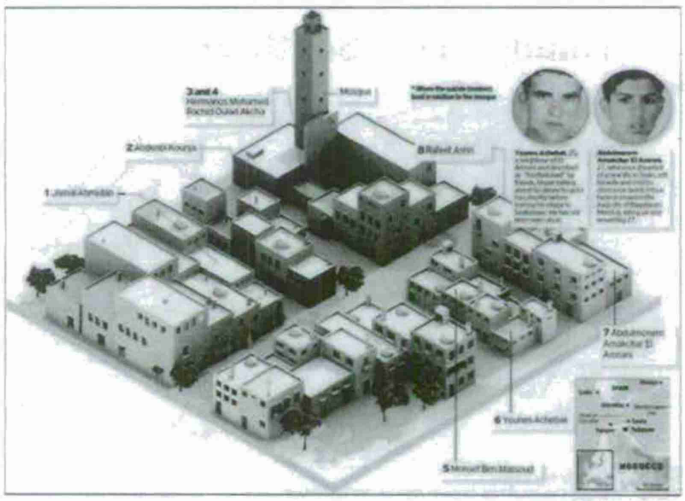


Fig. 2 Physical proximity of members of a Moroccan Cell.

core values – family, country, religion – are typically absolute and inviolate. Note that aspects of the above attributes drive US actions in Afghanistan, which may clash with analogous, but different aspects of a Middle Easterner's culture.

D. Anatomy of Terrorist Networks

A considerable effort has been spent to create the "John Jay Artis Transnational Terrorist " database (JJATT) on the evolution and structure of terrorist networks (Atran et al 2008). Key issues are the motivation for joining a network and conditions that favor an individual finding and joining cells aimed at terrorist activities. Unlike non-violent groups, the need for trust among members is critical. The Madrid and Moroccan units provide classic examples: physical proximity and friendship as starters to ensure trust, the desire to do something "significant", and the feeling that one's community's interest are being ignored or subject to injustices. High school friends, team-mates, neighbors in an apartment complex (see Fig 2) and kin are examples that lead to the strongest bonds. (Note consanguineous marriage is much higher in Muslim nations, hence there trust is easier to ensure over the potential cell members.) Exceptions to the above are when a cell matures and need other resources (i.e. explosives or money.) Then outreach increases risk, and the augmented cell typically becomes more fragile, with greater potential for fracture.

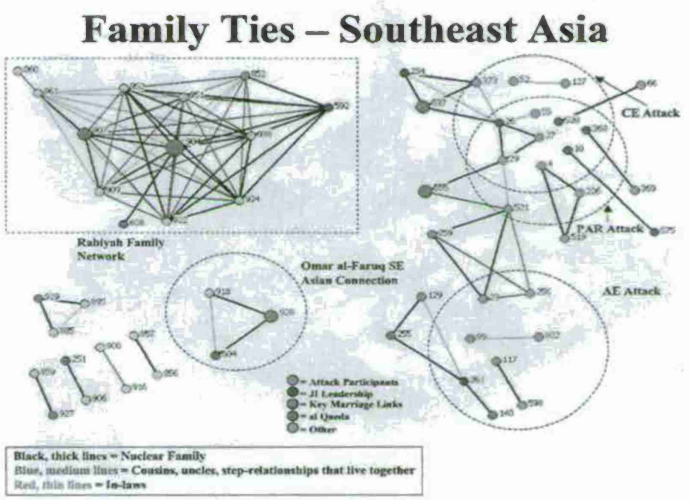


Fig. 3 A snapshot of part of a network showing the importance of family ties.

The same factors are also seen in the formation of non-violent networks (and also street gangs in the US), perhaps with different weights on risk or outreach. For example, over 50% of Facebook friends are typically within 50 mile radius. Exceptions are initiatives created via email, perhaps such as Tea Party, or educational initiatives aimed at a home country. (ASHA is a typical example where most US members have gone to the Indian Institute of Technology, and meet on a campus here.) This common background allows members to share similar goals and feel comfortable with each other. This is entirely analogous to membership in violent networks, where a key breeding ground is belonging to the same neighborhood or apartment complex.

• Perhaps a surprise is that the structure and evolution of a ***small*** terrorist network appear similar to the evolution of small non-violent ***social*** networks (Richards & Wormald, 2009.) (Note: the more global violent networks however, may differ from non-violent.)

This result is encouraging for those who study information exchange and belief revision in social networks. Furthermore, here is an opportunity to add priors to predict the complete structure of incomplete networks – i.e. to highlight possible missing links or nodes. Suggested priors are the relations among three measurement parameters to be describe below: Leadership, Bonding or clustering of individuals, and Diversity of members. The relation between these parameters is not arbitrary, as will be seen in section II-H. Such priors can be enforced at several network scales – but most effective will be on relations between neighbors and neighbors of neighbors for each vertex, where bonding indices should exceed the diversity index.

E. Multi-Agent studies: Colored Trails

The most popular platform for studying strategies used by agents or groups of agents in a negotiation game is probably colored trails. It was designed for use over networks, hence players can be at remote locations. Since its introduction in 2004 (Grosz et al) it is currently being used by over a dozen universities throughout the world. Members of the MURI have been key players in this development.

The game is quite simple, but designed to explore a range of strategies. Computer agents can assess and control the game play of human agents. Questions include: how do agents coordinate behaviors? How are groups

formed? What are the strategies in play? How best to change an agent's strategy?

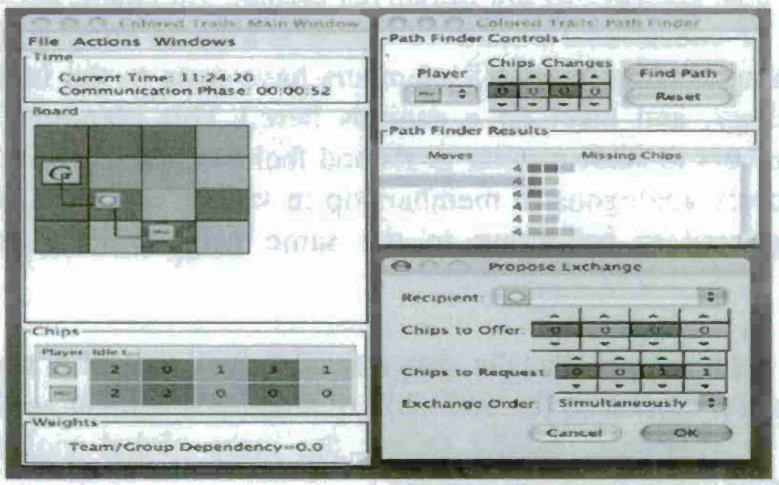


Fig 4 The Colored Trails interface. Chips are at lower left; game board above chips

The rules of the game are easy to grasp: there is a rectangular board of colored squares (see Fig 4). Each player is given a starting position and a goal position, and a set of colored chips. Players advance to their goal by moving to adjacent squares. But such moves are allowed only if the player has a chip of the same color as the square. If not, the player must negotiate a chip exchange with another player. The scoring for the game depends on the particular game played. In one simple version, how close a player is to the goal could be used, perhaps modified by the final number of chips held. The scoring function is varied to reflect different social policies and utility trade-offs.

Some of the recent findings are: (see references to Gal, Pfeffer, Ficici)

• Computer agents (and humans) that learn the social factors that influence human negotiation strategies can out-perform traditional game theoretic equilibrium strategies.

•Specifically, pay-off only representations for play lead to significantly less cooperation and benefit than is seen when negotiations based on social

factors take place, which seems to encourage helpful rather than selfish behaviors.

• A Nash equilibrium strategy leads to an equal or less social benefit than (WD) strategies that evaluate information (or actions) used to influence the behavior of opponents. (See section III-E for explanation of WD strategies.)

•The game, together with a computer agent, can assess and influence preferences, and thus can learn a human player's preferences (and decision strategies) given that the player is consistent.

• Computer agents that adapt their level of cooperation to the varying degree and helpful behavior exhibited by others will outperform computer agents that do not.

Note that several bullets are very relevant to negotiations involving sacred values. If offers during negotiations are simply competitive, then evaluating the degrees of cooperative behavior is absent and hence preferences will not be revealed during the game. Hence we have a corollary (which has been documented by experiment):

•Strategies that help reveal preferences during negotiations are more likely to lead to higher benefits for all parties.

II Model Frameworks

A. Narratives

Understanding cultural differences lies at the heart of how beliefs differ from one nation to another. Stories are an important window to this understanding. (Another approach, described later, is based on interviews that are then analyzed to determine similarities and differences among beliefs.) Stories prime youngsters on heroes and villains, on behaviors that are considered good or evil. A relevant example today is the different versions of the history of Israel given to kindergartners in Israel schools versus Palestinians in Gaza or West Bank. Each rendition frames the "facts" in a manner favorable to the modern interests and concerns. In a completely different domain, legal arguments can be seen as different stories that

support the defense on the one hand, or the prosecution on the other. (See report of Narrative Workshop: Richards, Finlayson & Winston, 2009.) These results are very relevant to how moral judgments are framed and played out in a culture.

Finlayson and Winston have initiated an effort to represent stories that allow a machine program to extract similarities (e.g. analogies) and differences among stories across cultures, or from news reports from completely different, sometimes conflicting sources (e.g. NY Times vs. Al-Jazeera.) This framework is based on the Story Workbench, which has now analyzed over two dozen folk tales from three cultures (Western, Russian and Chinese.)

• Higher level similarities among stories have been extracted using the Analogical Story Merging algorithm.

• These algorithms set the stage for automatic semantic analysis of different version of newscasts or media events, highlighting the differences in how the events are framed.

This same effort has also led to the formation of a world-wide group charged with creating a story database that will allow researchers to test the capabilities of their machine-based programs aimed at story understanding. (This is similar to the PennTree data bank used to evaluate natural language processing.) The first report of these efforts appears in Finlayson et al, 2010; a second report will appear in 2011 following a November 2010 meeting in Washington DC sponsored by the IEEE Artificial Intelligence Section.

B. Moral Decision-Making

Closely related to the above is decision-making. When encountering a new situation that presents a choice dilemma, analogies to past situations play an important role.

• Experiments on understanding the relation between cultural narratives and decisions invoking moral values support the hypothesis that cultural factors and analogies come into play (Dehghani et al 2009.)

This conclusion is based on stories that resonate with Iranian culture, but not the American culture. The examples involve sacrifice and are taken from Iranian folk tales. (Note these stories resemble the "Trolley problem" analyzed later in section III-F.) An example is:

"Pourya Vali was the most famous wrestler of his time. The morning before wrestling with a young athlete from another province, he goes to a mosque and sees the mother of the young athlete praying and saying " God, my son is going to wrestle with Pourya Vali. Please watch over him and help him win the match so he can use the prize money to buy a house". Pourya Vali thinks to himself that the young wrestler needs the money more than he does, and also winning the match will break the heart of the old mother. He has two choices: He can either win the match and keep his status as the best wrestler in the world or, he could lose the match and make the old mother happy.
Even though he was known not to ever lose a match, he loses that one on purpose. "

After reading the base story that presents the culturally accepted decision for resolving the dilemma, an analogical version of the story is given to participants (a group of 364 Iranians and another group of 48 Americans.) who were asked for their choice. Eighty-five percent of the Iranians chose the solution presented in the base story, whereas only forty-five percent of the Americans made that choice (i.e. a chance result for the two options.) See Dehghani et al 2010 for more details and other experiments.

• The key point is that decisions involving (sacrificial) moral choices are tied to a culture, unlike strictly utilitarian or economic decisions.

To model these results, and others, Forbus's group at Northwestern have developed a new framework and algorithms based on a combination of qualitative modeling and analogical processing. The languages used in qualitative modeling provide significant advantages over traditional numerical models because they directly express the causal theories commonly used in cognitive decision-making. (See section II-G for similar arguments.) Qualitative models also avoid the over-commitments that numerical models can entail, both in terms of amounts and precision of numerical data required to make predictions. Examples include:

• A model that represents ideas for blame assignment that can include the attribution of blame to multiple agents. [Tomai & Forbus, 2007, 2008].

• Demonstration that analogical generalization could be used to construct cultural models out of qualitative representations of interview data, which, for example, could be used to recognize group membership (Dehghani et al 2007). See also an analysis of food web resources for Menominee and competing majority American groups in Wisconsin (Dehghani et al 2007).

• A qualitative order of magnitude representation that incorporates sacred/protected values into a computational model of moral decision-making, including capturing factors that vary quantity sensitivity (Dehghani et al 2008).

An important distinction between sacred and secular (utilitarian) values is that the former are focused on acts (behaviors) not outcomes. In some contexts (or cultures) a higher weight should be given to the analogical solution over utilitarian choices, especially if the decision involves a sacred value. Some of these findings have been incorporated into MoralDM, a computational model suitable for analyzing moral decisions.

• MoralDM offers both analogies and rule-based principles to determine culturally sensitive decisions in situations involving both utilitarian (secular) and sacred values (Dehghani et al 2009). This model has been validated against 12 scenarios using the Ritov & Baron (1999) examples. The analogical component played an important role.

Most recently, Northwestern has augmented their advanced Explanation Agent Natural Language Understanding system (EA NLU) to handle counterfactuals, logical and numerical quantification, and quotation (Tomai & Forbus, 2009). This system has application to the semi-automatic generation of formal representations for stories, such as those in the MoralDM case library (See also section II-A.)

C. Cultural Signatures

Ethnographic research suggests that human cultures possess signatures – traits that might be described as coherent across a domain of action and beliefs. Such signatures allow actors within cultures to anticipate responses

across a wide range of contexts. (See section III-D on Tacit Beliefs.) Work in this area has been advanced by Bednar and Page, with the introduction and analysis of two interacting traits:

A culture with most individuals exhibiting the same set of traits is deemed ***coherent.***

A culture with traits that are related to one another are said to be ***consistent.***

Note that one might expect traits to vary considerably and bear no relation to one another. However, typically this is not the case. For example, if a player cheats are cards, there is no surprise if he also cheats on taxes, or more clearly if he runs red lights. For these cases the behaviors are consistent.

• A consistency-conformity model for cultural formation is proposed: Social pressures to conform are balanced against the need for traits (or beliefs) to be consistent with other traits.

The model represents each agent by a vector of M attributes having one of A values. The attributes represent behaviors, dispositions, customs, etc. If two attributes take the same value A, then the agent is consistent over these attributes. Overtime, these attributes and values may change, subject to external pressures. The relevant pressure is in a social context where the mean attribute values for a group differ from that of an individual. For the individual to conform, he must revise the attribute values, which may break consistency. The authors explore time to reach equilibrium under various scenarios.

If pressures to conform or to be consistent are varied, different levels of diversity results. For example, in a society where the relative tendency to conform is high relative to the consistency tendency, people may be less consistent but more similar.

This result impacts how surveys are conducted. Depending on the type of questions asked, a culture/country can appear either more or less heterogeneous. For example, if the survey questions ask about present behavior, a higher conforming society should appear less heterogeneous. However, if questions are hypothetical, a lack of consistency may give respondents a variety of possible behaviors, such in a moral context.

• The model offers a cautionary note: Depending on the questions asked in a survey, a less individualistic society like Japan could appear more heterogeneous than a highly individualistic society like the US.

Hence care must be taken to evaluate the structure or nature of a survey with respect to the cultural signature of the population being sampled. The results also impact the analysis of differences between and within political parties, or organizational interests, and why diverse societies may be better at problem solving (see Page 2007.)

D. Ritual and Tacit Knowledge

Tacit knowledge is motivated when an individual wants to conform to a social norm in a new context. This knowledge is not declarative, but inferred from observing the actions of others. Suppose you are in a line waiting to greet a foreign dignitary. Do you bow or so you shake hands? One possible strategy is to follow the actions of someone whose dress and mannerisms convey familiarity about the proper etiquette.

Tacit beliefs allow agents to acquire knowledge and infer accepted modes of behavior in new contexts without fully reasoning about all the consequences of their actions.

To begin to model the relation between tacit beliefs and actions, consider two binary actions denoted "+" (e.g. shake hands) and "-" (don't shake hands.) Let $K_i(R(+)$ indicate that agent I knows rule R+ is the norm, and $K_i(R-)$ means the action is against the norm. We can then identify four strategies:

S_1 (conservative) Choose action -. (Never shake hands.)

S_2 (norm abiding) Choose action + if $K_i(R+)$. (Shake hands only if you know it is the norm.)

S_3 (risk taker) Choose action + if ~~$K_i(R-)$~~. (Shake hands as long as you don't know it is against the norm.)

S_4 (reckless) Choose action + (always shake hands.)

This simple model leads to a set of knowledge predicates about the etiquette of "shaking hands", or any other binary norm. The full analysis takes several

other steps, one of the most important being how various agents interact with one another (See Gal et al, 2009.) These relations are captured by an ***interaction graph***. In the simple etiquette example, the interaction graph is a chain where only the behavior of one agent in front of another is considered. However, more generally interaction graphs can take many forms, such as a covered graph where one individual – the leader – is seen by every on else. Alternately the interaction graph can be bipartite, etc. A theorem is proven stating that whenever tacit knowledge is governed by underlying principles such as global consistency of knowledge, agents have enough information to "do the right thing."

An interesting corollary is related to ritual knowledge. Consider a group seated around a fire telling stories. The interaction graph is a ring, where all see one another. The sharing of expressions and reaction to the story add implicitly knowledge (about the group members). Social norms and aligned beliefs emerge very quickly. Thus narratives told and shared by a group typically invoke the power of rituals, especially when reinforced by an emotional factor (e.g. going to church is a very clear example.) Rituals are an important element to belief revision and the formation of strong beliefs that can lead to important social movements.

E. Entrenched Knowledge

When an entire community shares a narrative and that narrative is repeated and reinforced every so often by one and all, the underlying story becomes ***entrenched***. (See Doyle, Section III-D for formal definitions.)When a story is entrenched in a culture, it will provide a basis for why a decision should be considered rational. But, on other occasions, through overconfidence, the entrenched beliefs may lead to irrational decisions. The India-Pakistan relationship is replete with entrenched narratives of religion, race and other tacit beliefs. The two sides share a long and contentious history, and ways of thinking. Kasturirangan and Raghavan, 2009 show that on the Pakistani side, some entrenched beliefs included "we can fight a short

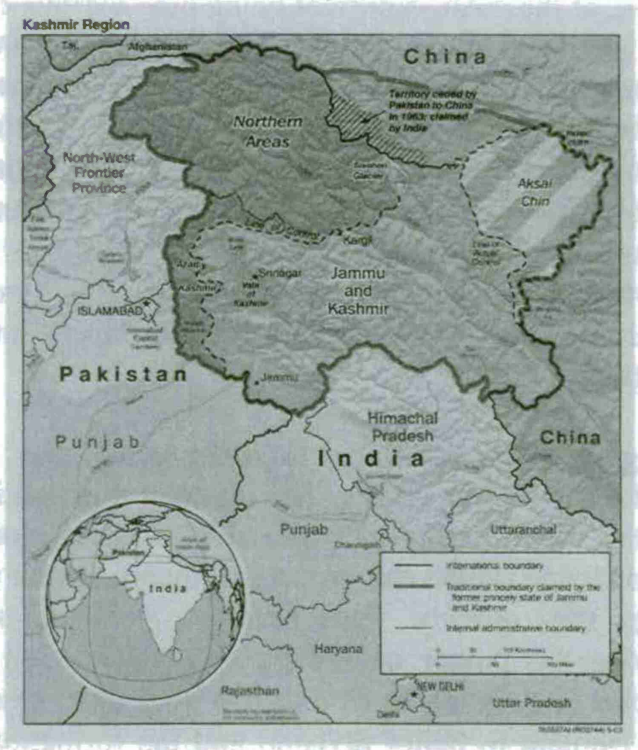


Fig 5. the Kashmir region provided the opportunity for two different MURI studies (Medin, Kasturirangan.)

sharp war", "the Hindu has no stomach for a fight". "the habit of submission and servility was engrained in the Indian people (due in part to the caste system)." The Pakistani military believed that one of their soldiers equal ten of more Indians". The belief that "the Hindu has no fight in him" drew on British, Orientalist conceptions of Indian society and history.

When the "Rann of Kutch" crisis emerged in 1965 in southern Kashmir, Pakistani forces attempted a "limited probe". India reacted only with mild enthusiasm and inadequate force, reinforcing the Pakistanis entrenched beliefs.

In retrospect, we see that the incursion did not resolve the Kashmir issues. However, in 1965 it is clear that entrenched beliefs about India played a key role in the Pakistani actions. In today's world, one might recast such beliefs as priors in a rational, Bayesian analysis. However, cultural stereotypes underlying actions are largely based on entrenched knowledge and decisions may entail a high degree of risk, especially if reliance is placed solely on a strict Bayesian analysis.

F. Case-Based Reasoning

Analogical reasoning is an important tool for individual decision-making, and several formal models are available (e.g. see sections II-A and II-B.) Surprisingly, however, little is know about the reasoning precedents in the public realm where proposals or positions may be subject to debate and where different precedents may compete or clash. A key element in such debates is how the problem at hand and the precedents are framed. Of special interest is how people of one culture can understand the framing process that people from another culture are likely to use (Atran, Axelrod, Davis, 2007; Atran and Axelrod, 2008.) Such an understanding embraces one's ability for influencing others.

• Axelrod, Atran and Davis propose a Theory of Case-Base Influence based on three concepts; salience, similarity, and compellingness.

Salience: The importance of a case irrespective of the current context.

Similarity: The degree to which one case is seen as providing a useful analogy to the other.

Compellingness: The extent to which a current problem is framed to be a compelling analog to a known case.

Given the above definitions, the Theory of Case-Based Influence proceeds with five premises:

1. Only the single most compelling case will be selected to frame the current problem.

2.The most compelling case will be the one with the highest product of salience and similarity.

3.The choice made in a particular situation will be the same (or analogous) choice that was made in the most compelling case, provided the outcome of that case was considered a success; if the outcome was considered a failure and the choice is binary, then the alternative choice will be made.

4.When a person tries to understand how others will frame a problem, the person will use his or her own beliefs about what cases the others find salient, and how the others will judge the similarity between of each of these cases and the current problem.

5.When a person tries to influence others, the person will advocate framing the current problem on the basis of two criteria. The person will choose among the cases that that he or she believes can be framed to the others in a manner that will support person's own choice, and from among these cases the person will advocate the one that he or she believes will be the most compelling to the others.

The scope of the Theory includes the consideration of "limited rationality", emotion, pseudo-historical cases (e.g. parables, religious examples), culture, and competition between policy advocates. To formalize the theory, several hypotheses are explicitly stated. One, for example, specifies how a person will choose among relevant cases, namely that their framing of the case will be deemed influential (hence dependent on that person's own view of the audience.) Clearly, then, an individual's own beliefs about the audience are a critical factor. (The same point was made in the Narrative workshop regarding factors relevant to story-telling.)

In support of the hypotheses, preliminary research has led to several observations, such as:

• People are heavily influenced by cases in which they are personally involved.

• The case used has high stakes and likely will still have emotional meaning for the individual.

• Cases will be seen similar to the current situation if the interacting governments are similar to those in the historical example.

Some of these points are considered, using the 1979 Iran hostage crisis or Truman's response to the invasion of South Korea in 1950.

Several suggestions are offered to improve negotiations, including those that involve sacred values. An illustrative example bears on cultural differences in the compellingness of cases:

• *The less a decision-making group (or an individual) relies on consequentialist reasoning, the more compelling will be cases involving quasi-historical or religious examples, parables, implied obligations, and*

inspiration. Conversely, the more a decision-making group (or an individual) relies on consequentialist reasoning, the more compelling will be historical cases.

The applications of this hypothesis include Arabs, Israeli settlers, Iranians, and American evangelicals when compared to decision makers from mainstream Western societies. For example, appeals by Al Qaeda's Osama Bin Laden or Iranian President Mahmoud Ahmedinejad are typically persuasive to their target audiences in these ways.

G. Multi-level cellular automata

Cellular automata have been popular in modeling agent interactions such as when agents may either cooperate of defect. Their advantage is two fold: (1) the rules for interaction are easy to implement in a language-like form and (2) the patterns generated as the interactions evolve give a visual picture of the state of the system.

• A simple conceptual advance over Wolfram's early studies (1983) is a multi-level cellular automata where quite distinct rules of engagement apply at different levels of the system. (Surprisingly, multi-level automata have been ignored.)

For example, in the social arena, individuals interact with one another to form groups having similar interests (or preferences.) On the next level, the groups may compete to gain control of the system as a whole, or to influence yet a third level that sets boundaries for group size and behaviors. Such a three level system reflects (a) the voter, (b) the parties and (c) the overarching federal government.

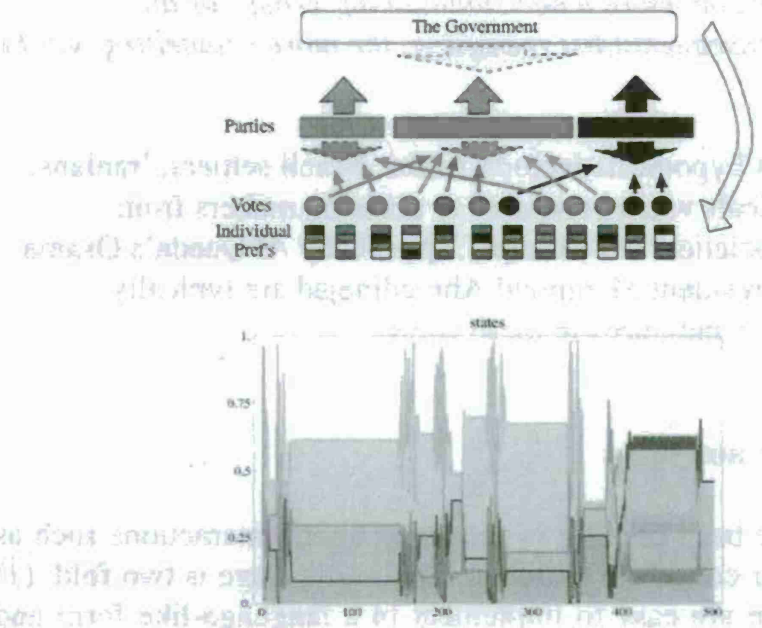


Fig 6. Top: The three level automata studied. Colored bars at bottom represent preference orderings for the agents. Lower: example of possible dynamic behavior. The colors indicate percent of the population favoring a preference.

• Three-level automata have very rich behaviors, but are easy to implement using language-like rules as constraints (see Koenderink & Richards, 2009)

Given the simplicity for implementation, it is easy to explore the equilibrium conditions for a system as complex as a population of voters who align beliefs to try to influence decisions at the governmental level. We can also model feedback systems, such as replicator dynamics, or history-based feedback (e.g. Shamma & Arslan 2005), and Lotka-Voterra population dynamics. An interesting property of multi-level cellular automata when applied to the socio-political realm is that phase transitions in state (e.g. group sizes) appear that mimic those seen in physical systems. This observation is potentially important for studying conditions for social unrest and upheaval.

H. Network Structure: Regularities

The structure of a network represented by nodes (e.g. individuals) and edges between nodes makes explicit the character of how information flows from one individual to another, or, alternatively, the similarity among a group of

individuals. For even 10 node networks, there are 10,000 varieties; for 12 nodes, over a billion. The challenge is to represent the network in such a manner such that different classes of networks can be identified, and, even more useful is to identify constraints on the different types of networks.

•The L B D representation shows that social networks have properties that are distinctive from other non-social networks. Furthermore, the representation permits a multi-scale view of a network, allowing for more detailed characterization.

The LBD representation is based on indices for leadership (L), bonding (B) and diversity of membership (D) , each of which is associated with a subgraph "motif" that captures an important aspect of social structures. The representation has the power of a scale space analysis which permits local as well as global views of the network structure.

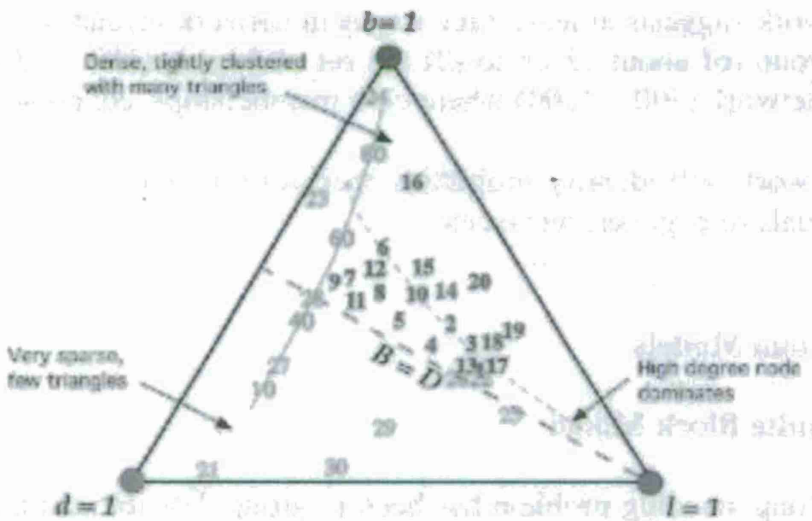


Figure 7 Illustration of the variability of LBD indices for different networks. For clarity, the raw LBD values have been projected onto the 111 plane (ie normalized by the LBD sum.) The red numbers refer to "social" networks; the green numbers are pseudo-trees or artificial (non-social) constructions. The dashed blue line indicates the range of Erdos-Renyi random graphs of probabilities indicated. (n= 100.)

Our analysis is based on 20 "social" networks where cognitive information is at play, as well as non-social networks such as Erdos-Renyi random graphs or the more passive non-cognitive spreading of disease or information flow, As mentioned earlier, we see no significant differences in

violent versus non-violent social network structures for networks less than 100 nodes.

In Fig 7, we show one of several regularities of "social networks", namely they all lie above the B = D divider (This is true even when the analysis is looking at local structures, such a nodes lying within two edge steps of the highlighted node.) *The implication of this regularity is that social networks tend to balance team bonding and the diversity of members, regardless of how dominant the leadership is*. (See Page 2007 for arguments.) Not shown are regularities associated with network size: for small networks having 100 nodes or less (roughly) there is one relation between L,B,D, whereas larger size networks have a different relation. This distinction is consistent with a suggestions by Dunbar, and curiously is roughly the point for the size of a face-book site – i.e. the rough maximum number of friends one individual will interact with at least on a semi-regular basis.

• The work suggests at least three stages in network formation: (1) the initial tight group (of about 12 or so (2) the set of friends (100 –150) and (3) the larger network (300 – 1000) where club memberships, etc come into play.

Future work will identify properties special to networks of 10,000 or more individuals (e.g. governments, etc.)

III Formal Models

A. Infinite Block Model

A long-standing problem has been to group data into similar categories. This ability is needed to analyze relations between individuals, their beliefs, their objectives and goals, etc., such as when an anthropologist gathers information about a community and its culture. For example, which groups of people have similar interests? What are these interests (especially shared beliefs)? How many distinctive groups are there? What cultural factors are underlying the groupings?

• The infinite block model, especially the CrossCat version, is one of the most advanced data classification methods available. (Kemp & Tenenbaum, 2008.)

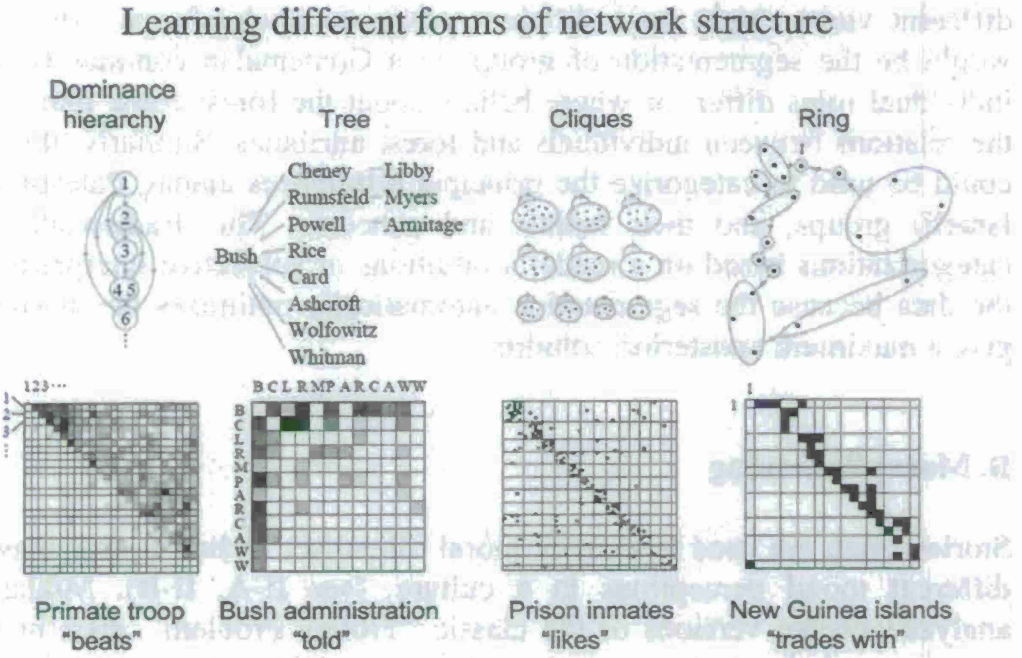


Fig. 8. The lower panels show "unscrambled" feature relations obtained by using the infinite block model algorithm. Above are the graphical forms derived from the block model outputs.

The data are typically a set of *n* unordered and ungrouped collection of elements (individuals, features, objects, etc.), with a binary value assigned to pairs indicating whether or not there is a relation between the two members of the pair. (This relation could be directed.) A simple visual depiction of these data would be an *n x n* grid (as in Fig 8), with elements of the grid either filled (black) or not. The array would look very random, unlike those above. The task is to shuffle the rows and columns of the grid until clustering is optimized, creating so-called "blocks" in the final array. The "blocks" reveal which elements belong together. These "blocks" can then be re-represented as a graph (upper portion of Fig 8.) CrossCat is a k-dimensional version, recovering multiple dimensions of groupings.

The latest Tenenbaum-Kemp model underlying the algorithm differs from earlier block models in the use of priors on the general form of the classes (e.g, perhaps hierarchical trees, perhaps bipartite, perhaps scale-free....etc.) This allow for sharper, more robust cuts through the data. Furthermore, multiple views of complex data sets become more interesting, because

different views might have different priors on model forms. An example would be the segmentation of groups in a Guatemalan community, where individual roles differ, or where beliefs about the forest come into play, or the relations between individuals and forest attributes. Similarly, the model could be used to categorize the principal differences among Palestinian (or Israeli) groups, and their beliefs and concerns. This framework avoids categorizations based on a reader's intuitions or subjective interpretation of the data because the segmentation automatically optimizes the grouping to give a maximum aposteriori solution.

B. Moral Reasoning

Stories often are used to convey moral dilemmas – choices that may have different moral perceptions in a culture. (see II-A, II-B). Mikhail has analyzed twelve versions of the classic "Trolley Problem", each of which presents a choice between an intentional homicide of one person to save many, or no action in which case many lives will be lost. These situations are analyzed from the viewpoint of an "intuitive lawyer" – one who possesses tacit knowledge of a rich variety of legal rules. (See section I-B for arguments that the law codifies moral or immoral acts in a culture.) Underlying the analysis are six deontic rules and several definitions, which underlie the formalization of a moral grammar.

• A key principle is the *Double Effect*. This principle distinguishes directly targeting individuals with harm from knowingly inflicting harm on individuals as an unavoidable by-product.

To illustrate, we summarize here four of twelve situations where an uncontrolled train is rushing toward five (unaware) men, who won't be able to get off the track in time. However, there is a side-track between the men and the train, and a switch can be thrown to divert the train. But in this case, one person, who also does not see the oncoming train, will be killed.

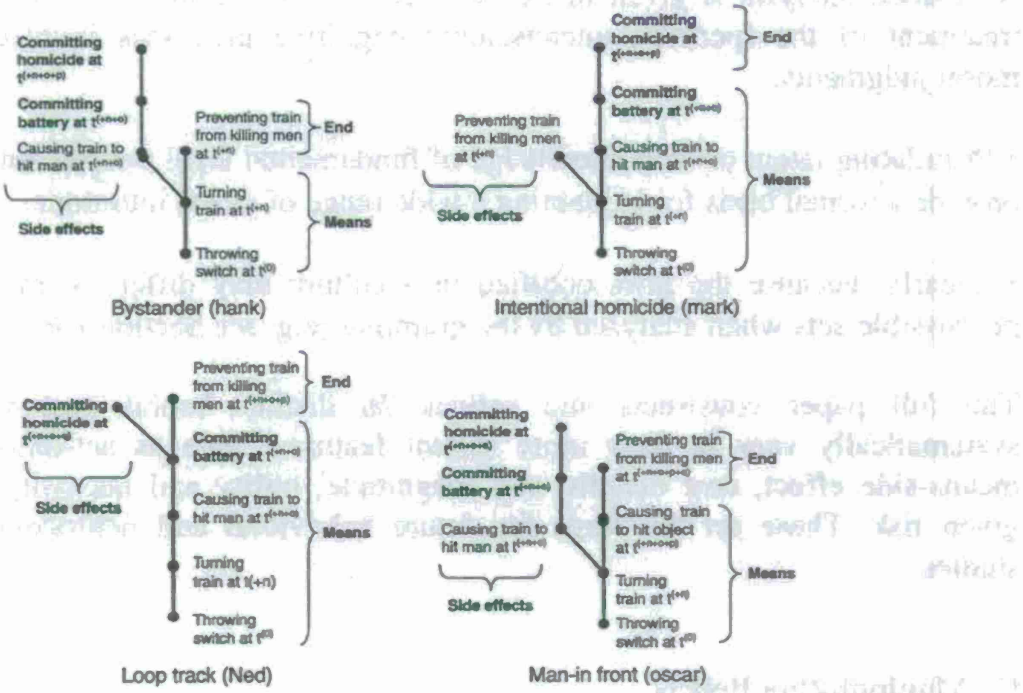


Fig. 9. Grammars capturing different version of the "Trolley Problem" (Mikhail, 2010.)

Four situations are illustrated in Fig 9. In the upper left, Hank is merely a bystander, and throwing the switch is a morally permissible act. A similar situation is in the upper right, but in this case, Mark knows and detests the man on the side-track. Throwing the switch is thus morally forbidden. In the loop track situation, throwing the switch will divert the train to a temporary side-track, which returns to the main line before the location of the five men. Throwing the switch will still lead to one man's death, but in the process the train will be slowed down, giving time for the five men to get out of danger. Again, Ned is morally forbidden from throwing the switch. Finally, on the side-track loop there is a man standing in front of a heavy object, which will slow the train and thus give the five men time to get out of the way when the train returns to the main line. Throwing the switch is permissible.

• The relation between the moral outcomes and the geometry (i.e. the grammar) in the above figure depends on the mean-ends relations, and the particulars of the side effects.

A detailed analysis is given in Mikhail, 2010. This is the most complete treatment of the (perhaps unconscious) cognitive processes involved in moral judgments.

• Postulating latent or tacit knowledge of fundamental legal conceptions can provide a formal basis for explaining a wide range of moral intuitions.

• Clearly, because the laws codified in a culture may differ, so may the permissible acts when analyzed by the grammar (e.g. see Section I-B)

The full paper constructs and refines 26 distinct moral probes that systematically vary in their more salient features, such as act-omission, means-side effect, cost benefit, and magnitude, utility and necessity of a given risk. These set the stage for future behavioral and neurocognitive studies.

C. Manipulating Beliefs

"Mechanism design" addresses how a social planner can design a framework to make policy decision based on messages from a population of agents. The aim is to derive rules that encourage agents to report truthful messages, and consequently lead to effective policy decisions for the social structure.

A number of studies have explored how beliefs (or bartering or information) propagate through a network. How can a social planner control such propagation? Mechanism design is a framework for formulating policy decision based on solicited messages from a population of agents. Rules are derived that give agents incentives to report truthful messages, thus ensuring effective policy decisions. Several issues emerge. These include the robustness of the policy framework, uncertainties in the network model linking agents, adversarial planners, and the underlying endogenous dynamics, especially when the environment is evolving.

These questions can be answered in part by methods of feedback control (Shamma & Arslan, 2007; Chasparis & Shamma, 2009). The approach is to augment an agent's decision at each time step by including the history of the probability distribution of possible strategies at that time step. In other words, the agent evaluates changes in the potential successes of the various

strategies. This inclusion introduces a dynamic that is akin to a myopic forecast that captures recent trends. Typically such forecasts would be used to stabilize outcomes to encourage convergence to an equilibrium. Here however, the goal is to move the system to a desired equilibrium – which is one of many possible. (See Pfeffer & Gal, section III-E for related studies.)

• Theorems establish stability and convergence properties under dynamic reinforcement.

• Not all Nash equilibrium may be stable (Some equilibria may be destabilized.) Hence there is an equilibrium selection device.

This work includes examples showing equilibrium selection for several 2x2 coordination games. The theme is extended to show how efficient networks can be found (i.e. the flow of benefits, for example) In these studies, solutions are found with agents having only local views of the network.

D. Belief Revision

Belief change lies at the heart of persuasion, negotiation and even how we revise frameworks for knowledge. Involved are desires, preferences, intentions, motives, habits, logic – to name a few relevant factors. The framework understudy by Doyle (2010) focuses on how logical and economic rationality is limited by structural or informational properties of the behavior.

•The novel insight in the formalization of belief change is a mapping to the concepts of mechanics.

Mechanics is based on forces that act on masses to change position (or distort the mass, etc.) The mass constitutes the portion of memory and configuration information that persists independently of motion. The (rough) mapping from mechanics to individual beliefs (or groups) is to allow types of individuals (i.e. different types of masses). The "state" of an individual (or group) is its location and velocity in a multi-dimensional space of properties such as preferences, desires, intentions, etc. Within this formulation one can invoke operations on states, for example, testing for

logical consistency and closure. A notion of coherence is defined by definitions of consistency and closure.

To capture the important distinction between sacred (or moral) beliefs and utilitarian (secular) beliefs, Doyle places partial orders on preferences for states. A simple example would be the reason $A \setminus\setminus B \mathbin{/\!/-} C \setminus\setminus D$ expresses the preference $\cancel{ABCD} < \cancel{ABCD} < \cancel{ABCD}$ for these three configurations. These orders can be augmented to include contexts.

Belief change is divided into two parts: motivated and accommodative. The motivated part specifies the properties required of the resultant belief state (Alchourr'on, Gardenfors and Makinson (AGM) which specify eight logical axioms that characterize the notion of revision – see table insert.); the accommodative part are changes needed to transform the motivated changes into an admissible mental configuration.

•An important step beyond the AGM formulation is the introduction of the notion of maximal preferred rational contraction. A definition of an informationally monotone preference order is given, which than can be shown to lead to a rational choice contract, consistent with the AGM axioms.

(–1) A – a is a theory whenever A is	(closure)
(–2) A – a ⊆ A	(inclusion)
(–3) If a / ∈ Cn(A), then A . – a = A	(vacuity)
(–4) If ⊬ a, then a / ∈ Cn(A . – a)	(success)
(–5) If ⊢ a ↔ b, then A . – a = A . – b	(equivalence)
(–6) A ⊆ Cn((A .– a) + a) whenever A is a theory	(recovery)

(–7) (A . – a) ∩ (A . – b) ⊆ A . – (a ◦ b) whenever A is a theory

(–8) If a / ∈ A . – (a ◦ b), then A . – (a ◦ b) ⊆ A . – a whenever A is a theory

These axioms mainly state fairly intuitive conditions, for example, that A . – a is always included in A (.–2); that contraction leaves A unchanged if a / ∈ Cn(A) (. –3); that the contraction omits a as long as . –4); that contractions by equivalent statements yield the same result (–5); and that adding a to the result of contracting by a

only yields conclusions present prior to the contraction (. –6).

Alchourr ´ on, Gardenfors, and Makinson also present eight axioms paralleling these that characterize the notion of revision, and show that the two sets of axioms are equivalent when connected by the Levi identity. They also show that these axioms are satisfied by a number of forms of belief revision.

Although the link between mechanics and logics may seem weak, it is useful to note that the Multicellular Automata framework proposed by Koenderink & Richards is equipped to handle exactly this type of information, and to explore equilibrium states. Note also that Doyle's approach complements that of Bednar and Page's treatment of conformity and diversity. Also, it provides a formal definition of entrenchment. Lastly, there a potential link to Pfeffer and Gal's proposal for "Wcll-Distinguished Strategies" that are based on informationally relevant choices.

E. Patterns of Strategic Reasoning

Negotiations or competitions require reasoning about another person's (agent's) likely decision. Clearly the reasoning process will depend upon the strategies in play. One obvious pattern for reasoning would be tit-for-tat, another would be to evaluate maximum expected utility, as in classical pay-off games. However, another quite different strategy (more like poker perhaps) would involve revealing information that might influence the opponent's behavior. For this last situation, it is possible to identify all possible reasoning patterns in a two person game. (Extensions to multi-agent play however can be made.)

Earlier, we provided evidence (from colored trail experiments) that greater benefits to plays occur when preferences are revealed by providing information to opponents. In the absence of revelation approaches, one would simply compute behaviors / actions based on various strategies such as "best response" or "Nash equilibrium" (See Page & Golman, 2009 for an analysis of equilibrium conditions in these and similar situations.) The

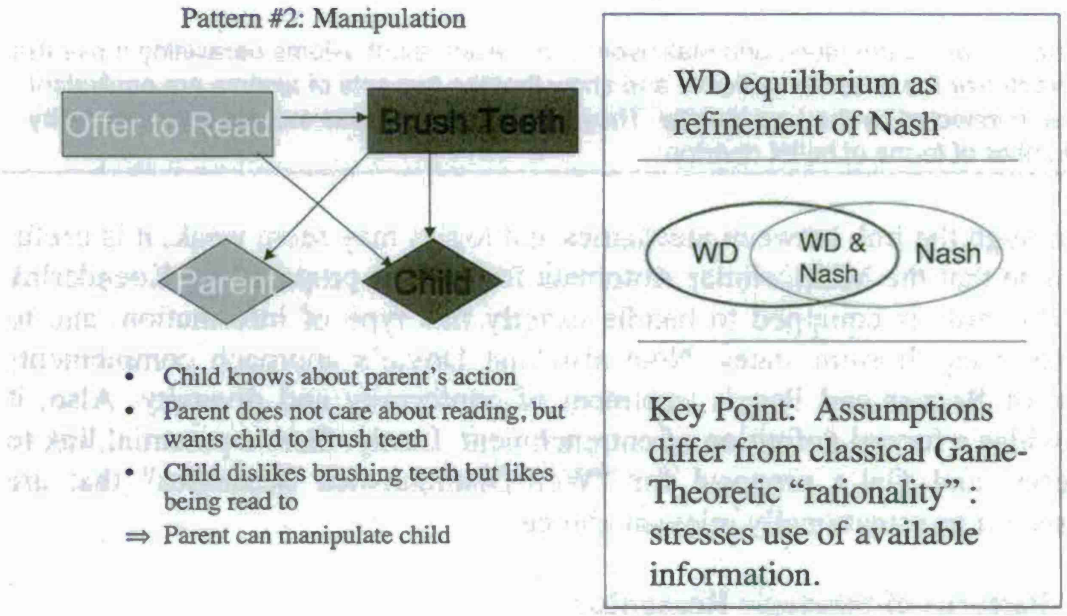


Fig. 10. The second of four possible reasoning patterns in a two-person game. (Pfeffer & Gal, 2008.)

alternate proposed here is that a player (agent1) understands that their actions can influence another's decisions (agent2) and thus will strive to choose actions that lead the opponent (agent 2) to react in the best way possible for the first player (agent1.) Surprisingly, if the second agent pays attention to the first agent's actions, then only four reasoning patterns need to be considered. Each characterizes a situation in which one agent cares about its decision and will use information available to help make the best decision. (The use of such information that distinguishes among actions is called a "well-distinguished" or "WD" strategy.) These patterns can be cast in terms of a multi-agent influence diagram (MAID.) In Fig. 10, a box is a decision D_i for an agent, a diamond is the utility or payoff U_i. In the more complex patterns, a circle is used to indicate variables such as additional information about the context that may or may not be revealed. The proof that these four patterns are complete is given in Pfeffer& Gal, 2008)

In multi-agent games where one of the players is the computer, it is possible to determine which of the four reasoning patterns (i.e. strategies) a player chooses and whether this is optimal given the programmed strategy of the computer agent. Experiments show that the WD strategies underlying the

reasoning patterns can outperform Nash (see Color Trails section.) Note that at the heart of this work is how the agent updates beliefs, given the actions of another.

A related effort is the representation of networks of influence diagrams (NID) for reasoning about agent's beliefs and decisions. These graphical structures are recursive, so that the mental model for one agent may contain mental models of other agents, such as in a multi-agent game. Each mental model is a NID is constructed using acyclic Bayes nets (Pearl, 1988) and MAIDs such as those shown in Fig 10. The formalism is powerful enough to use logic to collapse redundant edges in the NID, thus isolating the factors that constrain outcomes. An important theorem states that " Every Bayesian game can be represented by an equivalent NID whose size is linear in the size of the Bayesian game." Hence NIDs are more compact than Bayesian games in general. The most recent treatment is Gal & Pfeffer 2008.

IV Transitions

A. Briefings

Atran, S. "Devoted Actor versus Rational Actor Models for Understanding World Conflict," Briefing to the National Security Council, White House, Washington, DC, September 14, 2006, http://cstsp.aaas.org/content.html?contentid=621

Atran, S., Axelrod, R., Davis, R., *Terrorist Networks and Sacred Values*. Briefing to the National Security Council, White House, Washington, DC, March 2007; synopsis available at http://www.sitemaker.umich.edu/satran/files/synopsis_atran-sageman_nsc_brief_28_march_2007.pdf

Atran, S. *Terrorism and Radicalization: What Not to Do, What to Do.* Briefing presented to the U.S. State Dept/ UK House of Lords, November / December 2007; available at http://www.edge.org/3rd_culture/Atran07/index.html

Sageman, M. consultant to NYPD on terrorist networks and motivations. (2008)

Atran, S. "The Making of a Terrorist: A Need for Understanding from the Field", Testimony before the House Appropriations Subcommittee on Homeland Security, Washington, D.C. March 12, 2008, http://sitemaker.umich.edu/satran/files/atran_congress_12march08.pdf

Atran, S. "Pathways to and From Violent Extremism: The Case for Science-Based Field Research." Testimony before the Senate Armed Services Subcommittee on Emerging Threats & Capabilities, March 10, 2010; http://armed-services.senate.gov/statemnt/2010/03%20March/Atran%2003-10-10.pdf

Page, S. Testimony before US House Science and Technology Subcommittee on value of multiple models to understand complex social phenomena. 19 July 2010.

Note: In addition to the above, Scott Atran has given over two dozen public presentations at Universities and Symposia, and has written another two dozen popular press articles related to intelligent approaches to reducing terrorism and improving negotiations between cultures.

B. Software

Tenenbaum, J. B. Algorithms for relational forms (e.g. Infinite Block Model, CrossCat) of interest to several Gov't Labs (e.g. Sandia and Lincoln Labs) and universities (e.g. now available at CMU and Stanford through C. Kemp and T. Griffiths.) Further development is being pursued by Navia Systems for a scalable platform for government and public use.

Gal, Y et al's multi-agent Colored Trail platform for studying strategic reasoning is now used by a dozen universities in the US, Europe and Asia.

Finlayson. M. now is beta testing his Story Workbench for semi-automatic annotations of text, with interest indicated by several universities here in the US as well as in Europe. (This interest is in part due to our "Narrative Workshop" which has led to a team to pull together a story data bank.) The Java wordnet interface subroutine JWI has been released and over 4500 downloads have occurred http://projects.csail.mit.edu/jwi/.

C. DataBases

JJATT global terrorist database has over 2000 entries with individuals from over 20 countries. To help conduct visual and basic social-metric analyses if a software package call AUSTIN. <http://doitapps.jjay.cuny.edu/jjatt>

Story data base: this is in initiative involving several universities that begun in Apr. 2010. The aim is to create a public data base of annotated stories, similar to the PennTree data bank used for Natural Language study.

V. PUBLICATIONS: (Over 120 papers have appeared in refereed journals or refereed proceedings, or are under review. In addition there have been many presentations at national and international meetings, or at Government briefings.

Books

Atran, S. *Talking to the Enemy: The Dreams, Delusions and Science of Sacred Conflicts*, New York: HarperCollins (Ecco press), Paris: Editions Denoel, forthcoming.

Atran, S. & Medin, D. (2008). *The Native Mind and the Cultural Construction of Nature*. Cambridge, MA: MIT Press.

Page, S. E. (2010) *Diversity and Complexity*. Princeton University Press.

Papers

Atran, S. (January 2009) Who Becomes a Terrorist Today? *Perspectives on Terrorism*, 2 (5).
(http://www.artisresearch.com/articles/Atran_2008_Who_becomes_a_terrorist_today.pdf)

Atran, S. (2009) A Question of honor: Why the Taliban fight and what to do about it. (under review, based on our sacred values research).

Atran, S. (August 2008) "Band of Brothers": Civil society and the making of a terrorist. *The International Journal of Not-for-Profit Law*, 10 (4),.
(http://www.icnl.org/knowledge/ijnl/vol10iss4/art_3.htm)

Atran, S. Axelrod, R. & R. Davis (2007) Sacred Barriers to Conflict Resolution. Science 317, 1039-1040.

Atran, S. & Axelrod, R. (2008). Reframing sacred values. *Negotiation Journal, 24*(3), 221-246, available at, http://sitemaker.umich.edu/satran/files/negj0708.pdf

Atran, S. & Henrich, J. *The Evolution of Religion*. PNAS (Proceedings of the National Academy of Sciences), forthcoming.

Atran, S., Magouirk, J. & Ginges, J. (2008). The radical madrassah factor. *CTC Sentinel* (Combating Terrorism Center, West Point) 1(3) online, available at, http://www.ctc.usma.edu/sentinel/default.asp

Atran, S., Medin, D. and N. Ross (2005) Psychol. Rev. 112, 744-776.

Atran, S, Bennett, S. ,Fatica, A., Magouirk, Noricks, J.,Darcy, Sageman, M., and D. Wright. 2008. John Jay & ARTIS Transnational Terrorism (JJATT) Dataset. Retrieved from: <http://doitapps.jjay.cuny.edu/jjatt/>

Baker, C.L., Saxe, R., & Tenenbaum, J.B. (2009). Action Understanding as Inverse Planning. Cognition, 113, 329-349.

Bartels, D., Bauman, C. , Skitka, L.J. & D. L. Medin (Eds.), *Psychology of Learning and Motivation, Vol. 50: Moral Judgment and Decision Making*. San Diego, CA: Academic Press.

Bartels, D. M. & Medin, D. L. (2009). Are morally-motivated decision makers insensitive to the consequences of their choices? *Psychological Science*

Bednar, J. and Page, S. (2006) Can Game(s) Theory explain culture? The emergence of cultural behavior within Multiple games. (*Rationality and Society*, in press.)

Bednar, J., Bramson, A., Jones-Rooy ,A. & S. E. Page (2010)"Emergent Cultural Signatures and Persistent Diversity: A Model of Conformity and Consistency." *Rationality and Society*, in press.

Bednar,J., Chen, Y., Liu,X. & S Page. (2010) "Behavioral Spillovers and Cognitive Constraits in Multiple Games: An Experimental Study." Revise and resubmit to *Games and Economic Behavior*.

Bennis, W. M., Medin, D. L., & Bartels, D. M.(in press). The costs and benefits of calculation and moral rules. *Perspectives on Psychological Science*.

Broniatowski, D. and Kasturirangan, R. (2009). A Graph-Theoretic Analysis of Artifacts from Three Religions' Mystical Traditions. Submitted to The Journal of Cognition and Culture.

Chasparis, G. and Shamma, J.,(2009) Distributed Dynamic Reinforcement of Efficient Outcomes in Multiagent Coordination and Network Formation, (submitted for publication)

Chasparis, G. and Shamma, J., "Control of preferences in social networks under duopoly and network uncertainty." (work in progress)

Davis, R. (2010) Case-Based Reasoning: a Theory of Case-based Influence. Report of MURI subcontract to ARTIS, <rdavis@artisresearch.com>

Dehghani, M., Iliev, R., Sachdeva, S., Atran, S., Ginges J. & Medin, D. (2009) Emerging sacred values: The Iranian nuclear program (under review).

Dehghani, M., Forbus, K., Tomai, E., Klenk, M. (in press). An Integrated Reasoning Approach to Moral Decision-Making. In Anderson, M., Anderson, S (Ed.) *Machine Ethics*. Cambridge University Press

Doyle, J. (2010) Toward a quantitative theory of belief change: Structure, difficulty, and likelihood. NCSU Dept. Computer Science Tech Report #TR-2010-20. (on line)

Finlayson, M.A. Deriving Narrative Morphologies via Analogical Story Merging, in New Frontiers in Analogy Research (Proceedings of the 2nd International Conference on Analogy), edited by Boicho Kokinov, Keith Holyoak, and Dedre Gentner, New Bulgarian University Press, Sofia, pp. 127-136. http://www.mit.edu/~markaf/doc/finlayson.kokinov.nbupress.2009.137.pdf

Finlayson, M.A., Richards, W. and P. H. Winston (2010) Computational Models of Narrative: review of the workshop. AI Magazine, 97-100. Assoc. Adv. Art. Intell.

Fischhoff, B., Atran, S. & Sageman, M. (2008). M. Mutually Assured Support: A security doctrine for terrorist nuclear weapons threats. *Annals of the American Academy of Political and Social Science* (Special Issue: Terrorism Briefing for the New President) 618, 160-167, available at, http://sitemaker.umich.edu/satran/files/fischhoff_atran_sageman_published_mas.pdf

Gal, Y & A. Pfeffer Networks of Influence Diagrams: Reasoning About Agents' Beliefs and Decision-Making Processes Journal of Artificial Intelligence 33:109-147, 2008.

Ginges, J. & Atran, S. (in press). What motivates participation in violent political action: selective incentives or parochial altruism? *Annals of the New York Academy of Sciences*.

Ginges, J. & Atran, S. (in press). Flexibility and resistance to compromise over sacred values: An Indonesian case study, In D. Medin (ed.)

Ginges J., Atran, S., Medin, D. & Shikaki, K. (2007) Sacred Bounds on Rational Resolution of Violent Political Conflict. PNAS 104, 7357

Ginges, J. & Atran, S. (2008). Humiliation and the inertia effect: Implications for understanding violence and compromise in intractable intergroup conflicts. *Journal of Cognition and Culture*, 8, 281-294

Ginges, J. & Atran, S. (2009). Noninstrumental Reasoning over Sacred Values: An Indonesian Case Study. In D. Bartels, C. Bauman, L. Skitka, & D. Medin (eds.): *The Psychology of Learning and Motivation*. Vol 50, Chpt 6: 193-206.

Ginges, J., Hansen, I.G. & Norenzayan, A. (2009). Religion and popular support for suicide attacks. *Psychological Science*, 20, 224-230. (http://idisk.mac.com/jeremyginges1-Public/suicideattacks.pdf)

Haidt, J. & C. Joseph (2007) "The moral mind: how five sets of innate intuitions guide the development of many culture-specific virtues, and

perhaps even modules." In Carruthers, Laurence & Stitch (eds). "The Innate Mind: Foundations and Future" Vol 3. Oxford Univ. Press.

Haidt, Jonathan, Jesse Graham, and Craig M. Joseph (2008). "Above and Below Left-Right: Ideological Narratives and Moral Foundations." *Psychological Inquiry 19*(4) (special issue on political ideology), November 2008.

Hong, Lu & S. E. Page (2009) Interpreted and Generated Signals. Jrl Econ Theory 144, 2174-2196.

Iliev, R., Sachdeva, S., Bartels, D. M., Joseph, C. M., Suzuki, S. & Medin, D. L. (forthcoming). "Attending to Moral Values." In D. M. Bartels, C. W. Bauman, L. J. Skitka, & D. L. Medin (Eds.) *Moral Judgment and Decision Making: The Psychology of Learning and Motivation, Vol. 50*. San Diego: Elsevier.

Jiang, A.X., Leyton-Brown, K. & A. Pfeffer. Temporal Action-Graph Games:A New Representation for Dynamic Games. Uncertainty in ArtificialIntelligence (UAI), 2009. Available from http://www.cs.mcgill.ca/~uai2009/proceedings.html.

Joseph, Craig M. (in press). "Emotions, Intuitions, and Virtues: Ethics and Human (Second) Nature." In Verena Mayer and Mikko Salmela (Eds.), *Emotions, Ethics, and Authenticity*. Amsterdam: John Benjamins Publishing Company.

Joseph, C. & B. Riedel (2008) Islamic Schools, assimilation, and the concept of Muslim Amereican character. In: Being and Belonging: Muslims in the United States since 1911" K. Ewing (ed). Russell Sage Fdn Press, NY

Kamar, E., Gal, Y., & B. Grosz. Modeling User Perception of Interaction Opportunities for Effective Teamwork. Symposium on Social Intelligence and Networking. IEEE Conference on Social Computing. August 2009, Vancouver, British Columbia.

Kemp, C. & J. Tennenbaum (2008) The discovery of structural form. Proc. Nat. Acad. Sciences 105, 10687-10692.

Kasturirangan, R. and Raghavan, S. (2009). Tacit Beliefs and Entrenched Narratives: A Cognitive Study of India-Pakistan Relations. Submitted to Language and Cognition.

LeGuen, O., Iliev, R., Lois, X., Atran S., & Medin, D. (provisionally accepted) A Garden Experiment Revisited: Inter-generational Changes in the Sacred and the Profane in Petén, Guatemala. *Current Anthropology*

Magourik, J., Atran, S. & M. Sageman (2008) Connecting terrorist networks. Studies in Conflict and Terrorism, 31, 1-16.

Marden, J.R., Arslan, G. and J. S. Shamma (2009) Cooperative Control and Potential Games. III Trans Systems, Man & Cybernetics B; 39, 1393-1407.

Marden, J.R. and J. S. Shamma (2008) Revisiting Log-Linear Learning: Asynchrony, Completeness and Payoff-based Implementation. Submitted to *Games & Economic Behavior*

Mikhail, J. (2009) Moral Grammar and intuitive jurisprudence: a formal model of unconscious more and legal knowledge. In. B.H. Ross (Series ed) & Bartels, D.M., Bauman, C.W., Skitka, L. J., & D. I. Medin (Eds.) *Psychology of Learning and Motivation Vol 50: Moral Judgment and Decision Making*. Academic Press, San Diego.

Mikhail, J. (2010) Is the prohibition of homicide universal? Evidence from comparative criminal law. 75 Brooklyn. Law Rev. 493 - 508. (Symposium. Issue.)

Mikhail, J. (2010) 'Any animal whatever": Harmful battery and its elements as building blocks of human and non-human m oral cognition. (to be submitted)

Mikhail, J. (2010) Minding the Gap? Negligence and its elements as building blocks of human and nonhuman moal cognition. (to be submitted.)

Macindoe, O. (2010) Investigating the fine grained structure of networks. MS Thesis, MIT Dept. Elec. Engr. & Computer Science.

Page, S E., Sander, L. & C. Schneider-Mizell. (2007) Conformity and dissonance in generalized voter models. Jrl. Stat. Phys. 128, 1279 –1297.

Page, S. E. & R. Golman "General Blotto: Games of Allocative Strategic Mismatch" Public Choice Vol 138, pp 279 - 310, 2009.

Page, S.E. & & R. Golman (2009) "Basins of Attraction and Equilibrium Selection Under Different Learning Rules" Evolutionary Economics. 20, 72-75.

Page, S. E. & R. Golman (2009) "Individual and Cultural Learning in Stag Hunt Games With Multiple Actions" forthcoming in Journal of Economic Behavior and Organizations

Richards, W. (2005) Collective choice with uncertain domain models MIT-CSAIL tech report AIM-2005-054 (available on-line.)

Richards, W. Small Group Evolution. (MIT-CSAIL memo, Feb 08)

Richards, W., Finlayson, M.A. & Winston, P.H., Advancing Computational Models of Narrative. MIT CSAIL Technical Report No. 2009-063 http://hdl.handle.net/1721.1/50232

Richards, W. and O. Macindoe. Characteristics of Small Social Networks. MIT-CSAIL-TR-2010-033.

Sachdeva, S., Iliev, R., & Medin, D. L. (2009). Sinning Saints and Saintly Sinners: The Paradox of Moral Self-Regulation. *Psychological Science*, *20*(4), 523–528.

Shamma, J.S. and G. Arslan (2007) Dimensions of cooperative control. Chapt. 1 in Cooperative Control of Distributed Multi-Agent Systems. J Shamma (Ed). Wiley & Sons, Ltd.

Shweder, Richard A., Jonathan Haidt, Randall A. Horton, and Craig Joseph (2008). "The Cultural Psychology of the Emotions: Ancient and Renewed." In Michael Lewis, Jeannette M. Haviland-Jones, and Lisa Feldman Barrett (Eds.), *The Handbook of Emotions* (3rd edition). Guilford, CT: Guilford Press.

Tenenbaum, J. B., Griffiths, T. L. and Kemp, C. (2006) Theory based Baysian models of inductive learning and reasoning. Trends in Cognitive Sciences, 10(7) 682-687

Ullman, T.D., Baker, C.L., Macindoe, O., Evans, O., Goodman, N.D., & Tenenbaum, J.B. (2010). Help or hinder: Bayesian models of social goal inference. In Advances in Neural Information Processing Systems (Vol. 22). Cambridge, MA: MIT Press.

Conference Proceedings & Presentations (refereed)

Antos, D. & A. Pfeffer (2008) Identifying reasoning patterns in games. 23rd Conf on Uncertainty in AI.

Antos, A & A. Pfeffer. Using Reasoning Patterns to Help Humans Solve Complex Games. International Joint Conference on Artificial Intelligence (IJCAI), 2009. Available from http://ijcai.org/papers09/contents.php.

Baker, C.L., Goodman, N.D., & Tenenbaum, J.B. (2008). Theory-based Social Goal Inference. In Proceedings of the Thirtieth Annual Conference of the Cognitive Science Society (pp. 1447-1452).

Bartels, D. M. (2006) Morally-Motivated Judgment and Decision Making. Symposium at Meeting of Society for Judgment and Decision Making, Houston, TX.

Bartels D. M. & W. Bennis (2006) Deontology and Consequentialism in Morally-Motivated decision-making. Proc. Soc. for Judgment & Decision-Making, Houston, TX

Chasparis, G. and Shamma, J., "Efficient Network Formation by Distributed Reinforcement," 47th IEEE Conference on Decision and Control, Cancun, Mexico, Dec 8-11, 2008.

Chasparis, G.C. and J.S. Shamma (2010) Control of preferences in social networks. 49th Conf. on Decision and Control, Atlanta GA Dec 2010

Coen, M. H. (2005) Cross Modal Clustering. Proc. 20th Conf. On Artificial Intelligence (AAAI'05). (See also: MIT-Phd dissertation 2006: Multimodal dynamics.)

Dehghani, M., Iliev, R. & S. Kaufmann (2007) Effectsd of fact mutability in the interpretation of counterfactuals. Proc. 29th Annual Conference of the Cognitive Science Society, Nashville, TN

Dehghani, M. Unsworth, S., Lovett, A. & K. Forbus (2007) Capturing and categorizing mental models of food webs using QCM. 21st International Workshop on Qualitative Reasoning Workshop, Aberystwyth, UK

Dehghani, M., Tomai, E., Forbus, K., Iliev, R., Klenk, M. (2008a). MoralDM: A Computational Modal of Moral Decision-Making. *Proceedings of the 30th Annual Conference of the Cognitive Science Society (CogSci)*. Washington, D.C

Dehghani, M., Tomai, E., Forbus, K., Klenk, M. (2008b). An Integrated Reasoning Approach to Moral Decision-Making. In the *Proceedings of the Twenty-Third AAAI Conference on Artificial Intelligence (AAAI)*. Chicago, IL.

Dehghani, M., Tomai, E., Forbus, K., Klenk, M. (2008c). Order of Magnitude Reasoning in Modeling Moral Decision-Making. *International Workshop on Qualitative Reasoning (QR)*. Boulder, CO.

Dehghani, M., Iliev, R., Kaufmann, S. (2009). Causal Explanations in Counterfactual Reasoning. *Proceedings of the 31th Annual Conference of the Cognitive Science Society* (CogSci09).

Dehghani, M., Gentner, D., Forbus, K., Ekhtiari, H., Sachdeva, S. (2009). Analogy and Moral Decision Making. *Proceedings of the 2nd International Analogy Conference*. Sofia, Bulgaria.

Dehghani, M., Forbus, K. (2009). QCM: A QP-Based Concept Map System. In *Proceedings of the 23nd International Workshop on Qualitative Reasoning* (QR09).

Dehghani, M., Sachdeva, S., Ekhtiari, H., Gentner D., Forbus, K. (2009). The Role of Cultural Narratives in Moral Decision Making. In *Proceedings of the 31th Annual Conference of the Cognitive Science Society* (CogSci09)

Doyle, J. Cognitive Mechanics: Natural Intelligence Beyond Biology and Computation, Proceedings of the AAAI Fall Symposium on Naturally-Inspired Artificial Intelligence, Arlington, Virginia, November 7-9, 2008, pp. 35-37, AAAI Technical Report FS-08-06.

Ficici, S.G. and Pfeffer, A. (2008) Modeling how Humans Reason about Others with Partial Information, Seventh International Conference on Autonomous Agents and Multiagent Systens (AAMAS).

Ficici, S.G. and Pfeffer, A. (2008) Simultaneously Modeling Humans' Preferences and their Beliefs about Others' Preferences, Seventh International Conference on Autonomous Agents and Multiagent Systens (AAMAS).

Finlayson, M A and Winston, P. H. (2005) Intermediate features and informational-level constraint on analogical retrieval. In B. G. Bara, L. Barsalou, & M. Bucciarelli (edds.). Annual meeting of the cognitive science society (Vol. 27) Stressa, Italy.

Finlayson, M. A. (2008) Collecting Semantics in the Wild: the Story Workbench. Workshop on "Naturally Inspired AI", Wash. DC Nov 08; Tech Report FS-08-06, AAAI Press., Menlo Park CA

Finlayson 2009 "Deriving Narrative Morphologies via Analogical Story Merging" at the MIT Workshop on Computational Models of Narrative, Beverly, MA, 9 October 2009

Finlayson 2009 "Deriving Narrative Morphologies via Analogical Story Merging" at the Second International Conference on Analogy, Sofia, Bulgaria, 25 July 2009

Finlayson 2009 "Story Basis Sets and Intermediate Features" at the DARPA Workshop on EN-Mem Systems, Washington DC, 7 February 2009.

Forbus, K., Lockwood, K., Tomai, E., Dehghani, M., and Czyz, J. (2007). Machine Reading as a Cognitive Science Research Instrument. AAAI Spring Symposium on Machine Reading.

Forbus, K; Invited Keynote, 20th International Qualitative Reasoning Workshop, Dartmouth, NH. Title: Qualitative Representations as a Modeling Language for Cognitive Science, July 2006

Forbus, K. , Klenk, M. & T. Hinrichs. (2008) Companion Cognitive Systems: design goals and some lessons learned. Workshop on "Naturally Inspired AI", Wash. DC Nov 08; Tech Report FS-08-06, AAAI Press., Menlo Park CA

Gal, Y. & A. Pfeffer (2007) Modeling reciprocal behavior in human bilateral negotiation. National Conference on Artificial Intelligence (AAAI'07),Vancouver, BC.

Gal Y., Grosz B., Pfeffer A., Shieber S. (2007) The Influence of Task Contexts on the Decision-making of Humans and Computers Proceedings of the Sixth International Interdisciplinary Conference on Modeling and Using Context (CONTEXT'07), Roskilde University, Denmark.

Gal, Y., Grosz, B., Pfeffer, A. & S. Shieber (2006). Colored Trails: a formalism for investigating decision-making in strategic environments. Proc. Workshop of Reasoning, Representation, and Learning in Computer Games. 19th Int. Joint Conf. On Artificial Intelligence (IJCAI), Edinburgh, Scotland

Gal, Y., D'souza, S., Pasquier,P., Rahwan, I. & S. Abdallah. The Effects of Goal Revelation on Computer-Mediated Negotiation. Proceedings of the Annual Meeting of the Cognitive Science Society (CogSci), Amsterdam, the Netherlands, July 2009.

Gal, Y., Kasturirangen, R., Pfeffer, A. & W. Richards. How Tacit Knowledge Guides Action. Symposium on Social Intelligence and Networking. IEEE Conference on Social Computing. August 2009, Vancouver, British Columbia.

Gal, Y. a., Kasturirangan, R., Pfeffer, A., & W. Richards (2009). A Model of Tacit Knowledge and Action. IEEE Computer Society, Los Alamitos, CA, USA.

Ginges, J. (2006, November) Moral and Instrumental Values in Political Negotiation: Field Studies with Israelis and Palestinians. Paper presented at

the meeting of the Society for Judgment and Decision Making, Houston, TX.

Joseph, C. "Cultural Psychology and the Virtues: Some Theoretical Considerations and Empirical Findings." Paper presented at the 34th Annual Conference of the Association for Moral Education: "Faith, Democracy and Values: The Challenge of Moral Formation in Families, Schools, and Societies," University of Notre Dame, November 13 – 16, 2008.

Joseph, C. (2008) Islam, Virtual Theory and the Cognitive Science of Ethics. Ann. Mtg. Amer. Acad, of Religion, Chicao Il (1-3Nov08.)

Joseph, C. Images of the Ideal President and of Actual Presidential Candidates: Evidence from the 2008 Election.Paper presented at the 2008 Annual Meeting of the International Society for Political Psychology, Dublin, Ireland, July 14 – 17, 2009.

Kamar, E., Gal, Y., & B. Grosz. Modeling User Perception of Interaction Opportunities for Effective Teamwork. Symposium on Social Intelligence and Networking. IEEE Conference on Social Computing. August 2009, Vancouver, British Columbia.

Kemp, C., Tenenbaum, J. B., Griffiths, T.L., Yamada, T. and Ueda, N. (2006) Learning systems of concepts with an infinite relational model. Proc. 21st National Conference on Artiificial Intelligence (AAAI-06.)

Kemp, C. Goodman, N. D. & J. Tenenbaum (2007) Learning causal schemata. Proc. 29th Annual Conference of the Cognitive Science Society.

Kemp, C., Goodman, N. D., and Tenenbaum, J. B. (2008). Learning and using relational theories. *Advances in Neural Information Processing Systems 20*. [poster]

Kemp, C., Goodman, N. D., and Tenenbaum, J. B. (2008). Theory acquisition and the language of thought. *Proceedings of the Thirtieth Annual Conference of the Cognitive Science Society*. [poster]

Koenderink, J. J. & W. Richards (2009) Multi-level cellular automata and social dynamics. Proc. IEEE Int. Conf on Social Computing (SocialCom09),

Vancouver, B.C. Also presented at the Annual meeting of the Cog/ Sci Soc., Amsterdam.

Kotsalis, G. and J. S. Shamma (2010) Robust synthesis in mechanism design. 49th IEEE Conf. on Decision and Control, Atlanta, GA. Dec 2010

MacMahon, M. and Stankiewicz, B.J. (2006) Human and Automated Indoor Route Instruction Following. In Proceedings of the 28th Annual Conference of the Cognitive Science Society, Vancouver, BC.

Macindoe, O. and W. Richards (2010) Graph Comparison using fine structure analysis. Proc. IEEE Conf on Social Computing. #244.

Mansinghka, V. K., Kemp, C., Tenenbaum, J. B., and Griffiths, T. L. (2006).Structured priors for structure learning. Proceedings of the Twenty-Second Conference on Uncertainty in Artificial Intelligence (UAI 2006)

Pfeffer A., Gal Y., (2007) On Reasoning Processes in Games National Conference on Artificial Intelligence (AAAI'07),Vancouver, BC.

Richards, W. (2008) Modal Inference. AAAI Workshop on "Naturally Inspired AI", Wash. DC Nov 08; Tech Report FS-08-06, AAAI Press., Menlo Park CA

Richards, W. & N. Wormald (2009) A representation for small group evolution. Proc. IEEE Int. Conf. on Social Computing (SocialCom09). Vancouver, B.C.

Richards, W. and O. Macindoe. (2010) Decomposing Social Networks. Proc. IEEE Conf on Social Computing #205. Minn. MN.

Sachdeva, S. and Medin, D.L. (2009), Group Identity Salience in Sacred Value Based Cultural Conflict: An Examination of Hindu-Muslim Identities in the Kashmir and Babri Mosque Issues. *Proceedings of the 31st Annual Meeting of the Cognitive Science Society*, CogSci2009. Hillsdale, NJ: Earlbaum.

Shafto, P. Kemp, C., Mansignhka, V., Gordon, M., and Tenenbaum, J. B. (2006). Learning cross-cutting systems of categories. Proceedings of the Twenty-Eighth Annual Conference of the Cognitive Science Society.

Tomai, E. & K. Forbus (2007) Plenty of blame to go around: a qualitative approach to attribution of moral responsibility. In Proc. Qualitative Reasoning Workshop, Aberystwyth, UK

Tomai, E. & K. Forbus (2007) Narrative presentation and meaning. AAAI Fall symposium on Intelligent Narrataive Technologies, Washington D.C.

Tomai, E., Forbus, K. (2008). Using Qualitative Reasoning for the Attribution of Moral Responsibility. *Proceedings of the 30th Annual Conference of the Cognitive Science Society (CogSci)*. Washington, D.C.

Tomai, E. and Forbus, K. (2009). EA NLU: Practical Language Understanding for Cognitive Modeling. *Proceedings of the 22nd International Florida Artificial Intelligence Research Society Conference*. Sanibel Island, Florida.

For a complete listing of publications, see: http://groups.csail.mit.edu/belief-dynamics

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
25-10-2010	Final Performance Report	01-06-2005; 31-8-2010

4. TITLE AND SUBTITLE

Computational Models for Belief Revision, Group Decision-Making and Cultural Shifts

5a. CONTRACT NUMBER
FA9550-05-1-0321

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Whitman Richards (P.I.)

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Massachusetts Institute of Technology
Computer Science and Artificial Intelligence Lab
32-364
77 Massachusetts Ave,
Cambridge, MA 02139

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFOSR/RSL
Suite 325
875 N. Randolph St.
Arlington, VA 22203

10. SPONSOR/MONITOR'S ACRONYM(S)
AFOSR/RSL

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
AFRL-OSR-VA-TR-2012-0557

12. DISTRIBUTION / AVAILABILITY STATEMENT

Public release

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This MURI (1) explores how beliefs support and lead to certain actions in one culture but not another, and (2) develops computational models that further our understanding of the relation between beliefs, decisions, and actions. One key requirement for such models is to distinguish the different roles played by instrumental or secular values versus sacred values, which differ widely across cultures. The goal of these models is to provide formal explanations for how the beliefs of individuals affect group and individual actions, and how groups with shared interests evolve. Such models are an important step toward understanding and predicting the dynamics and actions of groups. They are fundamental to an understanding of how actions of a group may be altered by belief revision, by either internal or external pressures (including force). They are also needed for strategic reasoning in negotiations, where beliefs in different cultures may lead to what appears to be irrational proposals, yet are seen as rational in that culture.

15. SUBJECT TERMS
Computational Models, Decision-Making, Negotiations, Sacred Values, Social Networks, Terrorist Motivations, Strategic Reasoning, Cultural Differences

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
Unclassified					Whitman Richards
a. REPORT	b. ABSTRACT	c. THIS PAGE	UU	50	19b. TELEPHONE NUMBER *(include area code)* 617-253-5776